

早稲田大学大学院 理工学研究科

博士論文審査報告書

論 文 題 目

唇動作と音声による
多重バイオメトリック個人認証に関する研究

Multimodal biometric person
authentication
using lip movement and voice

申 請 者

市野	将嗣
Masatsugu	Ichino

情報・ネットワーク専攻

情報ネットワークシステム研究

2008 年 2 月

家庭や会社から PC のみならず種々の端末機器がネットワークを介して相互に接続でき、また誰でも安心して簡単に利用できるネットワークサービスを実現するうえで、本人の認証は不可欠な対策の一つである。この個人認証の手段には、知識や所有に基づく方法と本人固有の特徴を用いたバイオメトリック個人認証がある。ここで、バイオメトリック個人認証は、忘却、紛失などの危険性が少なく、かつ認証精度の個人に依存する要因が少ないことから最近注目されるようになっており、研究開発とともに標準化検討が活発に進められている。しかしながら、単一の生体情報だけを用いるバイオメトリック認証では、利用環境あるいは負傷、疾病などにより適用できない場合や十分な認証精度が得られないという問題が生じる。このため、複数の生体情報を用いる多重バイオメトリック個人認証が実用化に向けての現実的な解として認識されており、具体的な適用に関する研究開発がなされつつある。

本研究では以上の背景のもとで、唇動作と音声を用いた多重バイオメトリック個人認証に着目している。唇動作と音声は発話動作の連続性を考慮した認証を行うことができ、認証精度の向上を図る手段として有望である。また、唇動作と音声は発話動作から同時に取得することが可能であるため、ユーザへの負担が少なく、さらにそれぞれの生体情報について独立に照合した結果を統合することにより、実用上問題となる照明変動や雑音などに起因する認証精度の低下を抑える対策として利用できる。さらに、発話動作から得られる唇動作と音声は同期した情報であることから、なりすましを困難にする効果も期待できる。

本論文では、唇動作と音声から得られる生体情報の連続性と非線形性に着目して、多重バイオメトリック個人認証の精度向上を可能とする手法を提案しており、以下、各章毎の概要を述べ、評価を加える。

第 1 章は序論であり、本研究の目的と背景を明らかにし、本研究の位置付けと概要を示している。

第 2 章「話者認識の従来手法と問題点」では、本研究で着目している唇動作、音声を用いた従来のバイオメトリック個人認証方法に関して調査した結果をまとめるとともに、抽出した課題に関して検討を加えている。そして、これらの考察から、唇動作および音声認証アルゴリズムに核非線形相互部分空間法を用い、多重バイオメトリックスの統合アルゴリズムに核非線形判別分析を用いることを提案している。

第 3 章「非線形性を考慮した唇動作個人認証」では、発話時における唇動作に対して核非線形相互部分空間法を用いたテキスト指定型個人認証方式を提案している。

唇の動きを用いた認証において、唇動作データが非線形構造を持つ場合には、Eigenlip のようにデータの線形性を仮定したアルゴリズムでは識別面を正確に近似することが出来ず、認証精度の低下を引き起こす結果となる。そこで、まず扱う対象である唇動作データの分布と軌跡について調べ、これらは非線形性を示す可能性があることを実験的に確認し

ている。また、同一音声の唇形状は似通っていることから、識別面も非線形構造を持つことが予測される。以上の考察を踏まえ、非線形アルゴリズムの一つである核非線形相互部分空間法を適用することで、線形のアルゴリズムである相互部分空間法や **Eigenlip** 法、また区分線形識別器である混合ガウス分布モデルと比較して安定かつ高い認証精度が得られることを示した。

第 4 章「核非線形相互部分空間法の処理量削減手法」では、唇動作や音声認証アルゴリズムとして用いる核非線形相互部分空間法の処理量の減少方法を提案している。

核非線形相互部分空間法では、データを部分空間で表現するのに固有値問題を解く必要があるため、処理量はサンプル数の 3 乗に比例する。また、核非線形主成分分析で求められた基底同士の類似度は、全ての学習サンプルと入力サンプルの間での核関数を計算することにより求められるため、線形のアルゴリズムである相互部分空間法に比べると極めて処理量が多い。そこで本章ではサンプルを削減することで処理量を減少する方法を検討している。すなわち、核非線形相互部分空間法の処理量減少のためのサンプル削減手法として、ランダムサンプリング、一定サンプリング、**Tipping** が提案したスパースカーネル PCA、**Smola** が提案したスパースカーネル PCA、**Condensing**、**Chunking** またクラスタリングとして **K**-平均法を用いて、核非線形相互部分空間法と組み合わせたときの各サンプル削減手法の比較実験を行った。その結果、クラスタリングがサンプル削減手法として優れていることを示した。

第 5 章「非線形性を考慮した音声個人認証」では、核非線形相互部分空間法を用いた音声に基づく話者認証アルゴリズムを提案し、実験的に有効性を示している。

音声による話者認証においては、一般的に混合ガウス分布モデルが用いられているが、音声データの分布を考慮したアルゴリズムの妥当性の検証は必ずしも十分に行われていない。例えば、データの分布に湾曲構造のような非線形構造が認められる場合には、カーネルマシンに代表される非線形アルゴリズムを用いることが有効である可能性がある。そこで、扱う対象である音声データの分布について調べ、これらが非線形性を示すことを実験的に確認している。そのうえで、連続的な入力音声を積極的に利用して核非線形相互部分空間法を適用することを提案し、混合ガウス分布モデルと比較して高い認証精度が得られることを示した。

第 6 章「唇動作と音声の多重バイオメトリクスによる個人認証」では、唇動作と音声の認証器から出力されるスコアを統合するアルゴリズムに核非線形判別分析を用いたテキスト指定型個人認証方式を提案している。

多重バイオメトリクスを用いた認証において、さまざまなアルゴリズムの適用が考えられるが、スコア空間において扱う識別面が非線形形状を持つ場合、非線形性を仮定したアルゴリズムを適用しないと認証精度が低下する要因となる。そこで、スコア空間について調べ、非線形性を

示す可能性があることを実験的に確認している。また、実用性を考慮すると、チューニングパラメータが少ないことや認証に要する時間が短いことが望ましい。これらの考察から、スコアの統合には核非線形判別分析を適用することを提案し、線形のアルゴリズムである線形判別分析や **AND**、**OR** また各認証器より出力されるスコアの平均値を用いる方法と比較して高い認証精度が得られることを示している。

第 7 章「正準角の統合による相互部分空間法」では、複数の正準角を用いる相互部分空間法を対象として認証精度を比較し、正準角の統合方法に関する指針を与えている。

相互部分空間法の認証精度向上に向けては、対象とするデータからどのように部分空間の基底を求めるかが主に議論されている。一方、相互部分空間法の類似度として複数の正準角を定義することができることから、複数の正準角を用いて認証を行うことが提案されており、それぞれの正準角には認証に有効な特徴を有する可能性があることが実験的に示されている。従来の正準角の統合方法として、複数の正準角の平均値を用いる方法や **AdaBoost** を用いる方法が提案されているが、これらの手法については比較検討がこれまで行われていない。従って、どの方法を用いて正準角を統合すればよいかを明らかにするため、唇動作データと音声データに核非線形相互部分空間法を適用させた認証実験において、複数正準角の平均値を用いる方法と **AdaBoost** を用いる方法を適用して比較した結果、**AdaBoost** が高い識別率を示すことを明らかにした。

第 8 章は結論であり、本論文の結論と今後の課題をまとめた。

以上要するに、本論文は話者認証の課題に着目し、発話から得られる唇動作と音声の非線形性と連続性を考慮した多重バイオメトリック個人認証アルゴリズムを提案するとともに、その精度向上と高速化の方法を明らかにしている。その結果、情報通信システムにおいてバイオメトリック認証を実用的な個人認証手段として適用するための知見を与えており、情報通信工学の発展に寄与するところが多い。

よって本論文は博士（工学）早稲田大学の学位論文として価値あるものと認める。

2008 年 2 月

審査員（主査）	早稲田大学教授	工学博士（早稲田大学）	小松 尚久
	早稲田大学教授	工学博士（早稲田大学）、 Ph.D.（スタンフォード大学）	松山 泰男
	早稲田大学教授	工学博士（早稲田大学）	小林 哲則